Available online at www.sciencedirect.com
ScienceDirect
Procedia CIRP 59 (2017) 172 – 177
www.elsevier.com/locate/procedia

The 5th International Conference on Through-life Engineering Services (TESConf 2016)

Using big-data and surface fitting to improve aircraft safety through the study of relationships and anomalies

Duncan Wooder[a], Alan Purvis[a,*], Richard McWilliam[b]

[a]School of Engineering and Computer Science, Durham University, Lower Mountjoy, South Road, Durham DH1 3LE, UK
[b]IBEX Innovations Ltd, Discovery 2 NETPark William Armstrong Way, Sedgefield, County Durham TS21 3FH

* Corresponding author. Tel.: +44 (0) 191 33 42437. *E-mail address:* alan.purvis@durham.ac.uk

Abstract

The aim of this paper is to assess the utility of a Big-Data approach to fault detection for 'systems of systems', utilising the derivation of empirical relationships identified through surface fitting. So-called Big-Data Integrated Vehicle Health Management systems do currently exist, but tend to analyse the health of vehicle systems based on the behaviour of individual sensors and readings. This paper proposes that it is possible to consider vehicle systems with a 'macro' approach and identify relationships between key variables which may not be initially apparent. Used in this paper is the open source flight simulation software FlightGear which has previously been assessed for the development of fault detection systems with positive results. The relationships found can be combined into a model of expected results against which real-time data is tested. Surface fitting and the assessment of 'goodness of fit' is used to identify these relationships. It is proposed that this technique need not be limited to fault detection in vehicle systems but is also applicable to other vital systems which require redundancy and constant health analysis. This paper concludes that this method is a viable approach and that relationships can be successfully identified for fault detection purposes.

Peer-review under responsibility of the scientific committee of the The 5th International Conference on Through-life Engineering Services (TESConf 2016)

Keywords: FlightGear; fault detection; big-data; integrated vehicle health management; surface fitting

1. Introduction

Commercial air travel is an increasily accessible form of transport and when considered on a fatalities per kilometre basis, proves to be the safest mainstream form of transport, at only 0.05 deaths per bn km. However, when compared in terms of fatalities per journey, travelling by air can be seen to be nearly 30 times more dangerous than bus travel at 117 deaths per bn journeys[1]. Of course, the average person does not take as many journeys by air as by bus, and, for this reason, the risk is often seen as reasonable. It could be said that aviation will always be unsafe, to some degree, due to the nature of the mode of travel. However, it is the ambition of many organisations, ranging from aircraft manufacturers to airline operators to reach an optimal level of safety in aviation, whilst maintaining the profitability of the form of travel. It is the aim of this project to develop theory in fault detection from Big-Data which could help secure improvements in both of these areas. Big-Data is the concept of the processing and analysis of extremely large, multi-variate data-sets, often polled at high frequencies.

It is consistently true that a large proportion of fatalities in aviation each year can be attributed, at least partially, to mechanical error[2]. Fault detection and prediction can be used to mitigate the risk to human lives caused by the potential for unexpected or undiscovered mechanical faults. It has been the opinion of Rolls Royce for some time that a large pool of data (Big-Data) can be used to both improve efficiency and safety[3] by finding or predicting the occurrence of faults. However, the detection of faults within complex engineering systems is a well-known challenge [4]. The engines which Rolls Royce produces are each equipped with hundreds of sensors which report in real-time to engineers based in the UK. During each flight, terrabytes of data are generated by the aircraft's engines. These data are analysed on-board, during the flight, and a distillation of these data is transmitted to the ground for maintenance action to either be scheduled in a few weeks time or for a ground crew to be dispatched immediately to the flight's destination. Upon landing, the entire data-set is available for download and analysis. Anomalies in pressure, temperatures and vibration measurements, amongst others, are investigated as potential indicators that an engine requires service. The expected values are generated from both simulated and experimental analysis. This paper suggests that a similar strategy of Big-Data analysis could be used on the aircraft's systems as a whole with similar aims.

Although proof of concept is provided by Rolls Royce, the complexity of modelling and analysing an entire aircraft's systems as compared to a single engine is far greater. It is noted

Peer-review under responsibility of the scientific committee of the The 5th International Conference on Through-life Engineering Services (TESConf 2016)
doi:10.1016/j.procir.2016.10.126

that an engine is subject to relatively constant conditions and has limited variance in input but an aircraft and its systems experience a vastly varying set of conditions due to environmental or consequential factors ranging from the weather and cruising altitude to the duration and route of a flight. An aircraft is, of course, also a much larger system than a single engine and, is, in reality, a 'system of systems'.

This paper assesses the viability of a Big-Data approach to fault detection and prediction in aircraft systems and uses the open-source flight simulation software FlightGear[5] to provide the modelling of a Boeing 737. Ben Morris (2015) demonstrates the utility of *FlightGear as a Tool for Real-Time Fault Detection and Self-Repair*[6] and proves that its extensive input/output (I/O) features allow for it to be used for this purpose. Boeing, a leading commercial aircraft manufacturer, uses their Aircraft Health Management system (AHM system), a form of Integrated Vehicle Health Management (IVHM) on-board aircraft to attempt to detect or predict faults in order to increase safety and reduce the financial and logistical impacts of failure. This system monitors individual components on a micro scale to detect or predict failure on an element-by-element basis. The data generated from individual components is compiled into a system report, but the system's health is not analysed on a truly macro scale. It is suggested by this paper that a possible approach to this macro analysis is to use Big-Datasets from 'healthy' flights to form a model of how an aircraft is expected to behave. Individual, and potentially unexpected, relationships may then be drawn from these data-sets and flight data compared in real-time in an attempt to assess the health of the aircraft as a whole. The aim of this paper is to assess whether Big-Data from an open-source simulator may be used as a platform for the development of such a system.

2. Theory

2.1. FlightGear

Key to the operation of FlightGear is the Property Tree. The Property Tree is described as the 'central nervous system' in documentation[5]. It features a hierarchical tree-like structure of low-level variables. These variables range in function from user interface to Flight Dynamic Model (FDM), the modelled mechanics behind the behaviour of the aircraft. There are numerous different methods of viewing and modifying these variables. The folders entitled *instrumentation* and *engines* include many of the variables which would be expected to be found in an avionics computer of a real aircraft. There are approximately 1000 variables in these folders and it is these upon which this paper focuses. The input/output capabilities of FlightGear are comprehensive and designed for implementing configurable testbeds for a variety of applications [7]. I/O is configurable from the program launcher, with the ability for the use of standard or custom protocols. The Generic Protocol[8] allows for the design of custom packets with chosen fields. It is possible to select between reading and writing from/to file, serial, User Datagram Protocol (UDP) or Transmission Control Protocol (TCP). Since, by default, the FDM updates at 120Hz, it is possible to import/output variables at this rate. The configuration of protocols is facilitated by an easy to read XML file, examples of which are provided. FlightGear includes the option for the creation of multiple I/O protocols to allow for the concurrent transmission/receipt of different data sets.

2.2. Function Fitting

MATLAB includes objects, *cfit* and *sfit*, for the creation and analysis of empirical polynomial curve and surface functions fitted to a data set. It is possible for the user to select between curve and line fits as well as choosing the *fitType* from a variety of options including linear, quadratic or higher polynomial curves/surfaces. MATLAB offers polynomial fitting up to 9th order terms for curves and 5th order in each direction for surfaces in 3-dimensional space. With surface fitting, it is possible to select different orders of approximation for the x and y dimensions.[9] MATLAB's *fit* function uses the linear least-squares method to minimise the sum of the squares of the error between each data point and the proposed curve or surface. Equation 1 below shows how the number of coefficients, k, scales with m (number of variables) and n (order)[10]. It must be noted that MATLAB's *fit* function does not operate beyond 3-dimensions. Instead, *lsqnonlin* may be used.

$$\text{Number of terms}, k = \binom{n+m-1}{m-1} \tag{1}$$

Equation 1 means that the expression for the number of coefficients needed to be found in a general case of n-dimensional surface fitting includes a factorial term with both the number of dimensions and the order of the desired polynomial. For this reason, the computation time for the finding of a least-squares fitted surface, which depends linearly on this number of coefficients to be found and linearly on the number of data points in the data set being fitted, will scale with factorial order.

The fit object includes the fit result (ie a list of function coefficients) as well as 'goodness of fit' statistics including the sum of squares due to error and the root mean squared error or deviation. A normalized root mean square error can be found by dividing the RMS error value by the difference between the maximum and minimum values of the dataset being fitted to (This method is referred to in this paper as Method A). However, in this case, it is also possible to normalise the RMS deviation value by the range in the z-direction of the fitted function over the x, y plane in consideration by replacing the maximum and minimum values with the maximum and minimum heights of the surface. Approximated values for these can be found using the decimation of the surface in two dimensions. (This method is referred to in this paper as Method B).

3. Methods

3.1. Extracting Data from FlightGear

Before extracting a data-set from FlightGear, it was necessary to plan a 'healthy' test flight. For this study, a flight between London Heathrow and Toulouse Blagnac in a Boeing 737-200 was selected. The trip distance was chosen to minimise the size of the resultant data-set but maximise the cruise time of the flight and the aircraft was chosen as it includes the most developed auto-pilot in FlightGear. The -200 variant of the 737 was selected as it is optimal for the trip distance. A